External and Internal Validity

Karina Davidson

Mount Sinai School of Medicine

Organization of Session

- Internal and External Validity definitions and CONSORT reporting elements
- G-study facets
- Cognitive maps of validity components

Internal Validity

The degree to which the chosen design establishes the cause-and-effect relationship between the treatment and observed outcome

AND

establishes the absence of a relationship implies an absence of a cause-and-effect relationship

Internal Validity

- So, for each possible design:
  - Does it rule out alternative explanations?
  - Will it convince intervention protagonists that the treatment isn't the active ingredient?

Threats to Internal Validity

- History
  - introduction of statins
- Maturation
  - spontaneous remission rates
- Testing
  - repeated exposure improves performance

Threats to Internal Validity

- Instrumentation
  - coder drift
  - different coder
- Regression
- Differential Selection
  - randomization failed at some level

Threats to Internal Validity

- Experimental mortality
  - less psychologically minded drop out
- Selection interactions
  - problematic randomization interacts with any of the other threats

Enhanced CONSORT guidelines

1 TITLE & ABSTRACT : How participants were allocated to interventions (e.g., random allocation", "randomized", or "randomly assigned").

INTRO

2 Background: Scientific background and explanation of rationale.

METHODS

3 Participants: Eligibility criteria for participants and the settings/ locations where the data were collected.

4 Interventions: Precise details of the interventions intended for each group and how and when they were actually administered.

5 Objectives: Specific objectives and hypotheses.

6 Outcomes: Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors).

7 Sample size: How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.

Randomization:

8 Sequence generation: Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification).

9 Allocation concealment: Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.

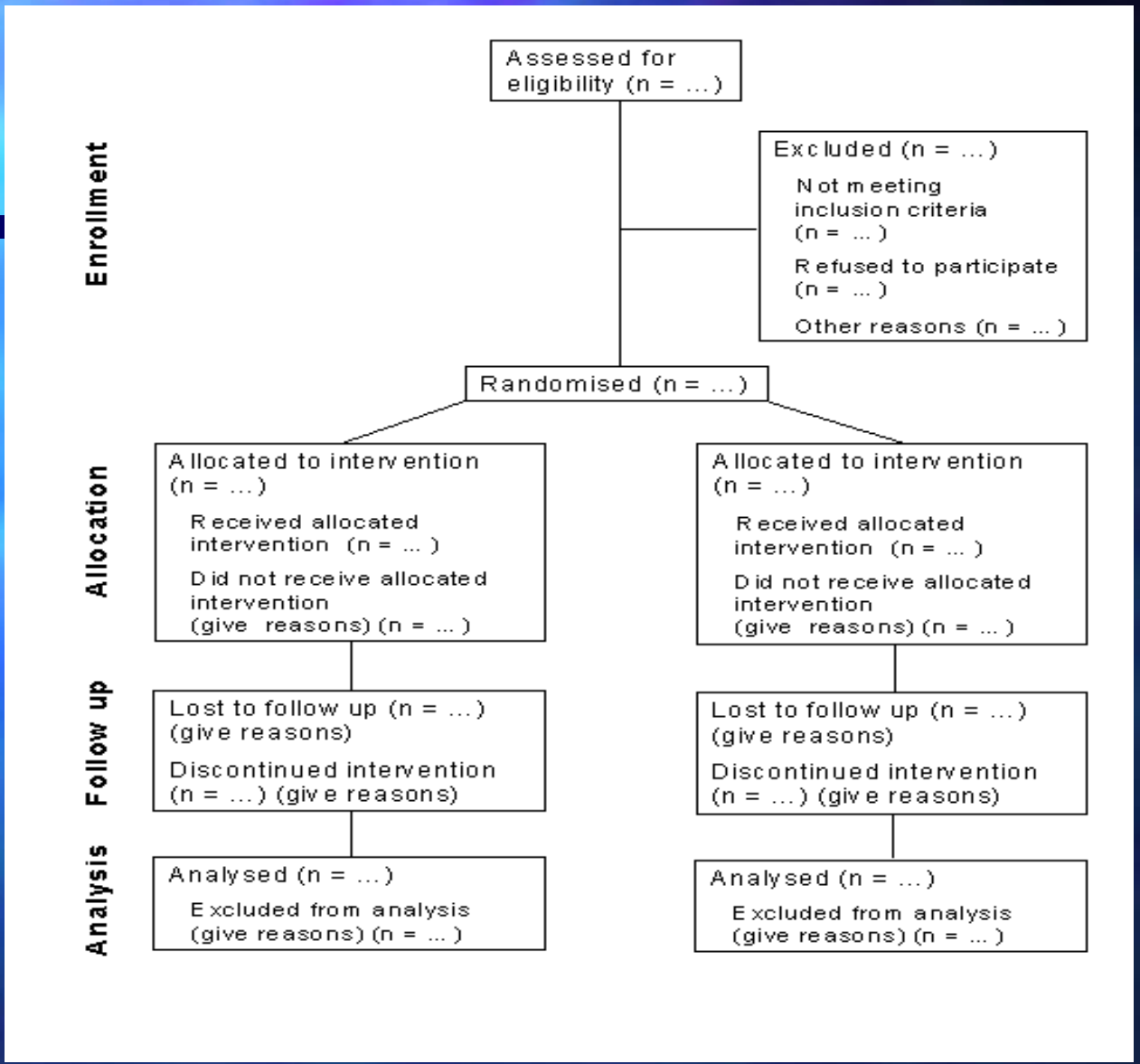
10 Implementation: Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.

11 Blinding (Masking): Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated.

12 Statistical methods: Statistical methods used to compare groups for primary outcome(s); Methods for additional analyses, such as subgroup analyses and adjusted analyses.

RESULTS

13 Participant flow: Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.


Assessed for eligibility (n = ...)
Excluded (n = ...)
Not meeting inclusion criteria (n = ...)
Refused to participate (n = ...)
Other reasons (n = ...)
Randomised (n = ...)
Enrollment
Allocation
Allocated to intervention (n = ...)
Received allocated intervention (n = ...)
Did not receive allocated intervention (give reasons) (n = ...)
Allocated to intervention (n = ...)
Received allocated intervention (n = ...)
Did not receive allocated intervention (give reasons) (n = ...)
Follow up
Lost to follow up (n = ...) (give reasons)
Discontinued intervention (n = ...) (give reasons)
Lost to follow up (n = ...) (give reasons)
Discontinued intervention (n = ...) (give reasons)
Analysis
Analysed (n = ...)
Excluded from analysis (give reasons) (n = ...)
Analysed (n = ...)
Excluded from analysis (give reasons) (n = ...)

14 Recruitment: Dates defining the periods of recruitment and follow-up.

15 Baseline data: Baseline demographic and clinical characteristics of each group.

16 Numbers analyzed: Number of participants (denominator) in each group included in each analysis and whether the analysis was by intention-to-treat". State the results in absolute numbers when feasible (e.g., 10/20, not 50%).

17 Outcomes and Estimation: For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (e.g., 95% confidence interval).

18 Ancillary analyses: Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.

19 Adverse events: All important adverse events or side effects in each intervention group.

DISCUSSION

20 Interpretation: Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.

21 Generalizability: Generalizability (external validity) of the trial findings.

22 Overall evidence: General interpretation of the results in the context of current evidence.

CONSORT Resources

- Resources: 1) Altman et al. AIM (2001, p. 663-694)
- 2) Winter 2001 Outlook submission on CONSORT

Behavioral CONSORT

23 A detailed report of the manner of testing, and success of, treatment delivery. (N/A okay)

24 A detailed report of the manner of testing, and success of, treatment receipt and enactment. (N/A okay)

25 Patient and therapist treatment allegiance or preference should be reported. (N/A okay)

26 Training of treatment provider(s) should be reported. (N/A okay)

External Validity

- Inclusion and exclusion criteria and characteristics of study patients

RE-AIM Criteria

Reach: State the population to which the study intended to generalize.

Reach: Report the number and rate of those excluded from the study (participation rates among those eligible).

Reach: Report the representativeness of participants by comparing study sample to known demographics of site/city/area. (N/A okay)

Efficacy: Report the intervention effects on quality of life.

Adoption: Report on methods for recruiting sites, as well as the inclusion and exclusion criteria used for enrolling sites. Provide rates of exclusion and refusal of sites.

Adoption: Report on the representativeness of settings studied by comparing to major characteristics of local area sites. (N/A okay)

Adoption: Describe the participation rate and characteristics of those delivering the intervention. (N/A okay)

Adoption: Describe how the treatment providers compare to intended eventual users of the intervention. (N/A okay)

Implementation: Report on the extent to which each component of the intervention is delivered as intended in the protocol. If different treatment providers delivered components ensure that this is clear in the report.(N/A okay)

Implementation: Report whether any specific measures of time OR costs are required to deliver the intervention.

Maintenance: If conducted in an applied setting, report on the extent to which the sites continued the intervention once the study was completed. (N/A okay)

Maintenance: Report on whether or not the study included at least a 6-month follow-up. (N/A okay)

Maintenance: Report in the flow diagram the participant rate of attrition (in addition to the participant number, as required by CONSORT).

Maintenance: If attrition rate is more than 10% in any study arm, provide analyses of the impact of attrition on outcome OR generalizability. (N/A okay)

G-Study

- Facets relevant to external validity
  - patients
    - gender
    - ethnicity
    - SES
  - therapists
  - sites
  - reimbursement schedules
  - etc

- Does the underlying theoretical model and presumptive active ingredient generalize to all facets????
- How does cognitive theory of depression do with this?